

Title: [What do we consider to be 'theth/ਠੇਠ/تھتھ' Punjabi? \[Warning: long post\]](#)

Author: me [user ThethPunjabi] (under an old Reddit account, now deleted)

Date: Sat May 21 01:31:31 2022 UTC

Sub-reddit: r/ThethPunjabi (I also recall republishing the same post on a variety of Punjab and Sikh-related subreddits, with minor altercations between them)

Post:

Hello, both this subreddit and wider interest in preserving the Punjabi-language has grown a lot and I think it would be worthwhile if we standardize, formalize, and define what we actually mean when we say 'theth' Punjabi in-regards to vocabulary. I appreciate the enthusiasm in the community about preserving Punjabi but it seems a lot of people are confused over vocabulary and what is authentic Punjabi and what is not. Therefore, I have decided to compile this post in-order to educate others on the linguistic nature of the Punjabi-language to increase your understanding of the language from an academic linguistic point-of-view. This may seem like a simple task at first but it might actually be more confusing and difficult to define what constitutes "authentic *theth* Punjabi" than you realize for a variety of reasons I will go into later.

What is *theth* Punjabi? I will try to give a definition for this linguistic concept. It is the purest form of the language that is most authentic to the Punjabi of the past. Today it is mostly spoken by elders and rural villagers due to the effects of language shift on the Punjabi youth, urbanites, middle and upper-class, educated-class, and the diaspora, which threatens the future vitality and distinctiveness of the Punjabi-language. Theth Punjabi has minimal influence (vocabulary-based, orthographic, and phonological influences) from other language of the Indian subcontinent and foreign ones. Theth Punjabi has been losing ground in recent-decades due to these influences caused by the mass-media, the Internet, Westernization, and globalization of our modern world and also due to stigmatization and discrimination targeting the Punjabi-language and its speakers.

First, I think it is helpful for us to classify the different kinds of vocabulary used in modern Punjabi. Thankfully, linguistics is an ancient field in the Indian subcontinent (Panini was born here!) so there has already been a ton of work done in this field so I will be borrowing concepts and categories already used by ancient and modern linguists of Indo-Aryan languages for Punjabi in my post. All Punjabi vocabulary can be classified into five (or six) groups that I will share below (these categories are also used for other Indo-Aryan languages like Hindi, Bengali, etc - I am merely appropriating their usage for Punjabi specifically in this post).

1) **Tadbhava** (ਤਦਭਵ/تدبھو)

Tadbhava (Sanskrit: तद्भव, IPA: [tədbʰəvə], lit. "arising from that") : is the Sanskrit word for one of the etymological classes defined by native grammarians of Middle Indo-Aryan languages (also applied to modern Indo-Aryan languages like Punjabi by contemporary linguists, grammarians, and etymologists). A "tadbhava" is a word with an Indo-Aryan origin (and thus ultimately descended and derived from Sanskrit) but which has evolved through language change in the Middle Indo-Aryan stage and eventually inherited into a modern Indo-Aryan language. In this sense, tadbhavas can be considered the native (inherited) vocabulary of modern Indo-Aryan languages and make up the core vocabulary of the languages.

Example: The core vocabulary of Punjabi would fall into this category, so most words you are familiar with can be used as an example. I will give you a specific example of a tadbhava word. The word for 'onion' in Punjabi is ਗੰਢਾ/كَنْدْهَا/Gaṇḥā. This word probably descends from the Sanskrit word सुकन्द/sukanda (which also means onion) but it has changed a lot from its ancestor and evolved naturally/inherited with the Punjabi language throughout all the Indo-Aryan language stages that it underwent throughout the ages (Old Indo-Aryan stage [Vedic Sanskrit to Classical Sanskrit], then Middle-Indo-Aryan stage [the local/regional Prakrit and then the Apabhraṃśa varieties that would become Old Punjabi], and then the New Indo-Aryan stage [Old Punjabi which would eventually become the modern Punjabi we know today]).

2) Tatsama (ਤਤਸਮ/تقسم)

Tatsama (Sanskrit: तत्सम IPA: [tətsemə], lit. 'same as that') : are Sanskrit loanwords in modern Indo-Aryan languages like Punjabi. They generally belong to a higher and more erudite register than common words, many of which are (in modern Indo-Aryan languages) directly inherited from Old Indo-Aryan (tadbhava). The words are borrowed in an unadulterated form from the predecessor language. The tatsama register can be compared to the use of loan words of Ancient Greek or Classical Latin origin in English (e.g. hubris). Tadbhavas are distinguished from 'tatsamas', a term applied to words borrowed from Sanskrit after the development of the Middle Indo-Aryan languages; tatsamas thus retain their Sanskrit form (at least in the orthographic form). This can be compared to the use of borrowed Classical Latin vocabulary in modern Romance languages.

In the modern context, the terms "tadbhava" and "tatsama" are applied to Sanskrit loanwords and descendant words not only in Indo-Aryan languages, but also in Dravidian, Munda and other language families and isolates of the Indian subcontinent.

Example: ਪੁਸਤਕ/پستک/pusataka (meaning: book) is a pure Sanskrit word that is totally unchanged from the original Sanskrit word (पुस्तक) and often used in Punjabi, therefore it is a tatsama. The proper Punjabi tadbhava word for 'book' is ਪੋਥੀ/پوٹھی/pōthī, probably descending from the same tatsama word (ਪੁਸਤਕ/pusataka from पुस्तक) showcased earlier.

3) Ardhatatsam/semi-tatsama (ਅੱਧਾ ਤਤਸਮ/ادھا تقسم)

These are semi-learned borrowings from Sanskrit. A word or other linguistic form borrowed from a classical language into a later one, but partly reshaped based on later sound changes or by analogy

with inherited words in the language. These words occur, for example, in the Romance and the Indo-Aryan languages.

Example: I actually cannot think of any. Can anyone else chime in here with examples for this category?

[author's edit: someone was kind enough to share some examples of ardha-tatsamas in Punjabi, I will share them below]

Some examples of ardh-tatsam words in Punjabi are:

ਈਸ਼ਰ Ishar from ईश्वर Ishvara

ਸੁਰਜ sUraj from सूर्य sUrya

ਸਿਮਰਨ simran from स्मरण smaraNa

4) **Deshaja/desya/desi** (ਦੇਸੀ/ديسی)

Deshaja/desya/desi (this is a very wide-encompassing category that can apply to many different kinds of words) : these words arise amongst the speakers themselves, without any relation to the predecessor language or originally borrowed from non-Indo-Aryan languages that are native to the subcontinent (such as Dravidian, Vedddoid, Munda [Austroasiatic], or Tibeto-Burman, language isolates, etc). Can also mean words that arose/coined/invented during post-Old-Indo-Aryan stages. It can also refer to indigenous words that we are unable to trace the origin of back to Old Indo-Aryan stages or to other indigenous language families of the subcontinent. These words tend to be very regionalized and dialectical. The difference between these words and tadbhavas is that tadbhavas are native vocabulary that was inherited/evolved from ultimately the Old Indo-Aryan stage (Sanskrit) whilst these words cannot be traced back to the Old Indo-Aryan stage. If a word was invented/coined in the Middle Indo-Aryan stage (for example, the Prakrit variety that would become Punjabi), it would be classed as a desi/desya word rather than the other categories. If someone comes up with a new Punjabi word today completely different from any other existing word, it would be classified as a deshaja/desya/desi word because it cannot be traced back to the Old Indo-Aryan stage.

Example: I actually cannot think of any. Can anyone else chime in here with examples for this category? There are examples in Hindi but I do not think it would be appropriate to share them here since our focus is solely on Punjabi.

5) **Videshaja/Vidēśī/foreign loanwords** (ਵਿਦੇਸ਼ੀ/وديشی)

Videshaja/Vidēśī (foreign-born) : these words are borrowed from other language families that are not native to the subcontinent and may or may not be in a modified form. Punjabi has a plethora of words from Persian, Arabic, English, etc. These can be modified or 'Punjabized' (perhaps due to having been borrowed originally a long time ago or for other reasons) when compared to the original foreign word that was borrowed, which could be its own category in itself (can call it a semi-foreign loanword/ardh-videsi).

So this category can be broken into two main groups: (1) historic foreign loanwords - that have marinated in Punjabi for centuries, and, (2) neo-foreign loanwords - words that have recently been adopted in recent decades from different sources (mostly English) and not underwent 'Punjabization/localization'

Example: ਕਿਤਾਬ/کتاب/Kitāba is a Punjabi word (meaning: book) that was borrowed centuries ago from from Persian کتاب (ketāb), which itself borrowed the word from Arabic كِتَاب (kitāb). First attested in Old Punjabi as ਕਤੇਬ/کتیب (kateba). It is considered a historic foreign loanword. Examples of neo-foreign loanwords is like when diaspora or urban Punjabis throw random English words into their Punjabi sentences.

Now that we have laid out the categories, we have to pick which ones are considered 'theth', this is more difficult than you realize. You may right away think that any foreign loanwords are not 'theth' (such as 'ਕਿਤਾਬ/کتاب/Kitāba' for *book*) but what are the alternatives? There used to be indigenous words with the same meaning (the tadbhava word 'ਪੋਥੀ/پوتھی/pōthī' comes to mind, which was derived from the Sanskrit word 'पुस्तक/pustaka', there is also the word 'ਗ੍ਰੰਥ/گرنه/granth' that originally referred to a book) but these indigenous words have lost their general definition and taken on a more rigid and specific definition since then and the foreign loanword has been adopted for the general definition (example is that 'ਪੋਥੀ/پوتھی/pōthī' and 'ਗ੍ਰੰਥ/گرنه/granth' now refer specifically to particular types of religious literature rather than their originally general meaning for any kind of literature, and kitab has since taken their place for the general definition). Therefore, it is fairly tricky. There are also words that are half-foreign and half-native. For example, the word for 'train' (as in railway travel) in Punjabi is 'ਰੇਲਗੱਡੀ/ریلگڈی/rēlagaḍī'. This word combines a foreign loanword (ਰੇਲ originally taken from the English word 'rail') with a native word (ਗੱਡੀ/گڈی), so is this a theth word? The reality is not as black-and-white as it seems and more grey. The only solution for these issues is inventing neologisms and repurposing indigenous words that have lost specific meanings overtime to regain them, or you can revive extinct words. Also, there may also be religious divides at play here. Punjabis who belong to the Islamic and Christian religions may feel more affinity to Punjabi words that were originally borrowed from Persian, Arabic, or European languages - this should be taken into account as well to prevent misunderstandings. It all depends how far you want to go down the path of linguistic purism.

My proposal for formulating a 'theth' Punjabi lexica-set and systematic rules of conduct is as follows:

Ideally, tadbhava words should be used with priority when writing or speaking in 'theth' Punjabi. If you cannot find an appropriate tadbhava or desi/desya word, you can then use a tatsama word but it is better if you 'Punjabize' it so it is more like an ardh-tatsama rather than a pure, unadulterated Sanskrit word being placed into modern-Punjabi. If this is still not appropriate, you can use a historic foreign loanword (from Perso-Arabic, Turkic, Mongolic, or other sources dating back to the mediaeval-period or earlier - you can also use older English loanwords that were adopted into Punjabi during the early-modern and colonial-era), historic loanwords are more ideal since they tend to be 'Punjabized' and

ardh-videsi. The last resort is using a straight-up neo-loanword like how Punjabis have started doing in recent years which is very un-ideal if we are trying to produce 'theth' Punjabi, so this must be avoided.

So the order of word preference is as follows for 'theth' Punjabi:

Tadbhava (1) > Desi/Desya (2) > Ardh-Tatsama (3) > Tatsama (4) [note - I place tatsamas as fourth but other Punjabis may change the order and prioritize ardh-videsi and historic foreign loanwords higher than tatsamas, this placement of tatsamas above them is my own preference as an Indian Punjabi, feel free to change up the order if you want for your own version of 'theth' Punjabi] > historic foreign loanwords and ardh-videsi (5) [note - some Punjabis may place this one third or even second due to personal preference, religious, or cultural reasons but I have placed it fifth in accordance with my own inclination] > neo-foreign loanwords (6) [note - I think this one should be avoided as much as possible, if that is not possible, neologisms can be invented, extinct words can be brought back, or if we must, the neo-loanword can be 'Punjabized']

What are your thoughts on this? I hope my post has been informative.

Disclaimer: I am not a trained linguist, this is just my past-time hobby as a member of the Punjabi community born in the diaspora. I simply have shared my knowledge from the various readings I have done over the years with you all.

I hope my post will be helpful for the ongoing language revitalization movement and struggle of the Punjabi-language.

Comments:

[\[-\]TimeParadox997](#) West Punjab | ਲਹਿੰਦਾ ਪੰਜਾਬ | 8 پنجاب ہند points 3 years ago*

Very interesting!

I think videshi (foreign loan words) should have subcategories.

There are a lot of words which have 'marinated' in Punjabi (and other Indo-Aryan languages) for a long time. Particularly words which came through Farsi centuries ago. This would be the 'ardh-videshi' (semi-foreign) category.

The other category of loan words would be the the relatively recent additions ('true' loan words?), particularly from English. Also Urdu (and Hindi).

Similar to this second category, in my opinion, are tatsam (direct Sanskrit loan) words.

Urduisation and Hindiaisation probably fall onto one of the latter two. And this (IMO) is a place where it can get tricky. If it's claimed that a word has come directly from Urdu (ie Farsi), we should know if that word actually did so, *or* was it in Punjabi from before? (Similarly with Hindi and Sanskritisation).

Not understanding how firmly or loosely rooted a word is *may* cause a person to 'hypercorrect' Punjabi.

<https://en.wikipedia.org/wiki/Hypercorrection>

[\[–\]](#)[deleted] 6 points 3 years ago*

Very good points. I agree that there are so many foreign loanwords in Punjabi that have been apart of the language for so many centuries that it would be odd to extinguish them completely from the language. These words need to be considered separately from random neo-loanwords (many examples from English could be used here).

We could separate them like this:

- 1) Historic loanwords (Perso-Arabic introduced during the mediaeval-period, European words introduced during the early-modern and colonial-period) - these words are very much apart of the Punjabi ethos and it seems very odd and artificial to separate them from the language for the sake of Indo-Aryan linguistic purism
- 2) Neo-loanwords (random words for new concepts and when people are just being lazy, are uneducated in Punjabi, or trying to sound “foreign”, examples being ‘internet’, ‘cellphone’, ‘school’, etc - usually always coming from English)

I have also noticed another phenomenon where tatsama words carry a higher sense of “prestige” in the minds of people and people like to replace the unique Punjabi tadbhava words with Sanskritic tatsamas. I am wary about this because I believe tadbhavas are the heart of modern Indo-Aryan languages and what makes them unique and distinguishable from other Indo-Aryan languages. This is more prevalent in Indian Punjab. I think in Pakistani Punjab, it’s more of a situation where Perso-Arabic historic (maybe new ones too) loanwords are replacing the tadbhava words.

[\[–\]TimeParadox997](#)West Punjab | ਲਹਿੰਦਾ ਪੰਜਾਬ | 5 | لہندا پنجاب | 5 points 3 years ago

I think in Pakistani Punjab, it’s more of a situation where Perso-Arabic historic (maybe new ones too) loanwords are replacing the tadbhava words.

As far as I know, it's not Persio-Arabisation, I think Urduisation might be more accurate to describe the situation

[\[–\]](#)[deleted] 5 points 3 years ago*

I've noticed people get confused on this subreddit with differentiating Punjabi and Hindustani (Hindi/Urdu) words. The truth is both languages are closely related so they share a lot of common vocabulary (cognates). However, usually there are some little quirks to tell them apart sometimes. For example, the word for "thanks" in Punjabi is dhanvaad, whilst in Hindi it's dhanYavaad, so a slight difference. Also, there are differences in pronunciation for various sounds (some sounds exist in Punjabi but do not exist in Hindustani, maybe vice-versa too) and of course Punjabi is tonal. As for historical loanwords, it is definitely tricky as sometimes you don't know if a Perso-Arabic word entered Punjabi or Hindustani first and if they have adapted since then differently in each language that adopted the word. These things need to be investigated on a case-by-case basis carefully.

[\[-\]Somebody890_4](#) 1 point 3 years ago

We say Mehrbani for thanks in Lehnda Panjab. Shukria is Urdu and Dhanyavaad is Hindi (Both Hindustani).

[\[-\]desi_miata](#) East Punjab | ਚੜ੍ਹਦਾ ਪੰਜਾਬ | 5 چڑھدا پنجاب points 3 years ago

Amazing post, very informative

You should post it on [r/punjabi](#) as well

[\[-\]](#)[deleted] 2 points 3 years ago

I am glad you liked it. I posted it there already!

[\[-\]Somebody890_4](#) 3 points 3 years ago

I think trying to remove all foreign words from Arabic, Persian, Turkish etc should not really be done because if this happens then Panjabi would become unintelligible and basically would be returned back to Prakrit and wouldn't even be Panjabi anymore. This would also cause us to lose our literature and music kind of like the Turks did when they Latinized their script. I think it is sufficient to try to purify Panjabi back to Modern Panjabi which dates back to the 12th Century. This would of course mean removing the Urdu/Hindi and English words that have entered Panjabi in recent years. We should also try to develop a system to develop new words like Computer, Refrigerator, Air conditioner etc like Arabic has done, so we wouldn't have to rely on the English words for these.

[\[-\]Super_Voice4820](#) Abroad | ਪਰਦੇਸ | 1 پردیس point 8 months ago

nah, jarnail is what many say for general, because these concepts have entered due to the british.

[\[-\]TimeParadox997](#) West Punjab | ਲਹਿੰਦਾ ਪੰਜਾਬ | 2 لہندا پنجاب points 3 years ago*

For me (as a Pakistani Punjabi), the order of preference would be as follows:

1. Tadbhava (natural evolution from Sanskrit)
2. (maybe) Desya (read definition in the OP)
3. Ardh-Videshi (historic foreign loanwords) and Ardh-Tatsama (evolved Sanskrit borrowings)
4. Tatsama (direct Sanskrit borrowings)
5. Videshi (neo foreign loanwords)

Note: I would have put Ardh-Tatsama below Ardh-Videshi, but the example of sūraj convinced me to raise it

Now, regardless of which order you like, we just need to make sure we know which category a word/phrase is in for *definite* before we make rash decisions on it's usage.

[\[-\]FaithlessnessHeavy75](#) 3 points 3 years ago

Um my opinion is.

That we should eradicate all foreign words and words from sanskrit instead develop a new vocabulary with counterparts for all those words or use Punjabiified version of them.

The reason why i am saying this is because Punjabis from west Punjab use General Punjabi + Middle eastern vocabulary, Punjabis from east Punjab use mixture of General Punjabi + Middel eastern and Punjabified Sanskrit volcabolary and Pahadis/dogras use General Hindi + Pure Sanskrit vocabulary. There was a time there was no such divide.

Which as you can see is used to divide us. So that we all sound different and then use it to justify that we all speak different languages. You will soon see, Saraikis will start using more Sindhi vocabulary intentionally and Hindkowans will start using more Pashto vocabulary, just so that they can sound apart and justify why they are not Punjabis.

We need to create new mechanics by which we can add new theth Punjabi words to the language and also preserve/ revive theth Punjabi vocabulary.

Basically a Sanjhi vocabulary.

[\[-\]TimeParadox997](#) West Punjab | ਲਹਿੰਦਾ ਪੰਜਾਬ | 5 | [پنجاب](#) | [لہندا پنجاب](#) points 3 years ago

That we should eradicate all foreign words and words from sanskrit instead **develop a new vocabulary with counterparts for all those words** or use Punjabified version of them.

I understand that you want to keep everyone united, but I fear that by developing a new vocabulary..., it would backfire and just create yet another group, rather than unite the existing groups.

[\[-\]](#)[deleted] 1 point 2 months ago

Panjabi language is not derived from Sanskrit.

[\[-\]](#)[Super_Voice4820](#)Abroad | ਪਰਦੇਸ਼ | 1 پردیس point 22 days ago

Confused ethnonationalist